

Complexity, boost symmetry, and firewalls

Ying Zhao

*Stanford Institute for Theoretical Physics and Department of Physics,
Stanford University, Stanford, CA 94305-4060, USA*

`zhaoying@stanford.edu`

Abstract

We find that the time dependence of holographic complexity is controlled by the Rindler boost symmetry across the horizon. From both the perspectives of collision energy experienced by an infalling object and quantum cloning, we see the breaking of this boost symmetry is closely related to firewalls, which in turn shows the connection between the time dependence of complexity and firewalls. We further identify the black and white hole interiors as two tapes storing different parts of the minimal circuit preparing the state. Depending on whether the quantum gates are being laid on the tape at a particular moment, each tape can be in two states: working, or locked. We interpret the existence of firewalls as the locking of tapes. We see that within this two-tape picture, various concepts: complexity, minimal circuit, quantum cloning, and firewalls fit together.

Contents

1	Introduction	1
2	Time dependence of complexity and boost symmetry	5
2.1	Known results about holographic complexity	5
2.2	Linear increase of complexity: Boost symmetry across the horizon	7
2.3	Breaking boost symmetry: Abnormal time dependence of complexity	10
2.4	More examples	15
3	Time dependence of complexity as a diagnosis of good interior	16
3.1	Boost symmetry breaking and firewalls	18
3.2	Boost symmetry breaking and quantum cloning: Firewall is necessary	19
4	Quantum circuit and firewalls	21
4.1	Future and past tapes	21
4.2	Two tape picture: Why are there both future and past horizons?	23
4.3	Tape locking: an interpretation of firewalls	27
5	Discussion	29
A	Complexity of perturbed thermofield double	31
A.1	Complexity of a precursor	31
A.2	Switchback from multiple precursors	32
A.3	Breaking spatial translation symmetry: Localized shocks	34
B	Collision energy and time dilation factors	35
B.1	Raychaudhuri equation	35
B.2	Landau-Lifshitz pseudotensor	37

1 Introduction

In [1] [2], complexity of a holographic state is conjectured to be dual to the volume of maximal surface anchored at boundary. In [3] [4], it is pointed out that the action in the Wheeler-DeWitt patch offers a better candidate for complexity. Both prescriptions give certain features of time dependence of complexity, like linear increase for sub-exponential time, switchback effect, e.t.c., which match quite well with what one expects from the

quantum circuit picture. In the first half of this paper, we explore the origin of these various features, and we'll see that the Rindler boost symmetry across a horizon plays a key role.

The picture of a tensor network lying in spacetime [5] [6] offers an intuitive explanation about the duality between boundary state complexity and bulk spacetime. It also makes explicit the idea of a black hole interior as emergent spacetime [7] [2] [4]. We can view the tensor network as a quantum circuit lying inside the horizon, recording the dynamics that creates the state starting from some simple state. The complexity is one important parameter characterizing this dynamics. In particular, if the state is made by inserting various precursors to the thermofield double, complexity keeps track of the dynamics of the CFT disturbed at various points.

On the other hand, so far we don't have a clear definition of complexity in continuum field theory. In the complexity-action prescription, we also need to pick a cutoff surface. In the first half of this paper we'll show that, starting from some mild assumptions, certain properties of the time dependence of complexity is robust despite these uncertainties.

In particular we will see that the time dependence of complexity is basically a consequence of boost symmetry across the horizon as indicated in Figure 1. Notice that the Schwarzschild time inside the horizon is synchronized with the Schwarzschild time outside. This boost symmetry is the origin of the linear time dependence of complexity. With the insertion of various precursors, the uniform dynamics is disrupted, and the bulk symmetry is broken. However, it's broken in a particularly simple pattern which results from the Rindler-like nature of near horizon geometry. We'll see that events near the horizon can play a crucial role. If they affect the synchronization, there will be abnormal time dependence of complexity, and vice versa. ¹

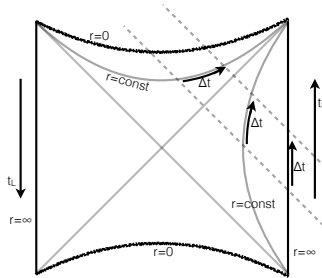


Figure 1: Boost symmetry across horizon. Schwarzschild time inside the horizon synchronizes with Schwarzschild time outside.

¹In this paper, we are talking about sub-exponential time. We don't really know what will happen after the complexity saturates.

We turn to the issue of firewalls in the second half of this paper. By calculations in general relativity, we'll see that the breaking of boost symmetry is directly related to the high energy collision experienced by an infalling object, so is a diagnosis of firewalls. From a different perspective, by exploring the operational meaning of this boost symmetry, we also see that a good symmetry is necessary for protecting an infalling observer from seeing quantum cloning. When this symmetry is badly broken, a firewall is then necessary to ensure nothing could enter the interior. This illustrates the connection between the time dependence of complexity and transparency of horizons as advocated by Susskind in [8].

By doing a more detailed matching between the picture of the quantum circuit preparing a state and the dual black hole geometry, we then identify the different roles played by future and past horizons. In the minimal circuit preparing the state, there can be both forward and backward Hamiltonian evolutions. They are stored in two different tapes. The future interior serves as the future tape storing the forward Hamiltonian evolution, while the past interior is the past tape storing backward Hamiltonian evolution. In this sense, the existence of past horizons means complexity can increase towards the past, and is a manifestation of time reversibility.

As a result, we give an interpretation of firewalls as the locking of tapes. When the quantum gates are being laid on a tape, we say this tape is working. Because tensor network supports spacetime, on the bulk side this means the corresponding interior is growing. When the interior grows, the boost symmetry is good, and if anything enters the interior, the quantum cloning is protected from being seen. At this moment the horizon is transparent. On the other hand, when no more quantum gates are being laid on a tape, we call this tape locked. On the bulk side, without a fast growing of interior, quantum cloning is no longer protected. A firewall instead will prevent anything from entering the interior.

The tensor network description of black hole interior is given by Hartman and Maldacena in [7]. The idea of a tape storing the quantum circuit behind the future horizon is also not new. It appears in [9] by Susskind. See section 3.3 and Figures 10, 17, 19 in [9]. The new ideas here are that, there are both future and the past tapes governing the evolution, and that firewalls are associated with the locking of one tape and serve as a protection for quantum cloning. We see that within this two-tape picture, various ideas: complexity, minimal circuit, quantum cloning, and firewalls fit together.

Here we should also emphasize the limit of this picture. The examples we studied are special. We know the minimal circuit preparing them, and their minimal circuits are made of Hamiltonian evolutions perturbed at various points. We don't know what will happen

for more general states. The lesson we learn here is, in order for the interior to be open to enter, it needs to be growing. On the quantum circuit side, there needs to be gates being laid on the corresponding tape.

For an exterior observer, a black hole is just a particle with certain properties [10]. What's particularly special about it is, there is spacetime behind the horizon, and an in-falling observer can enter. In this sense, the horizon is like a door to the interior region. We see that whether the door is open or not depends on the working status of the corresponding tape. You can only enter an interior when the corresponding tape is working. When the tape is not working, the entry is forbidden by firewalls.

In this paper we'll use the following assumptions:²

1. In holographic duality, the complexity of a boundary state is dual to some bulk geometric quantity inside the Wheeler-DeWitt patch.
2. Tensor networks support spacetime. They have locality down to AdS scale, and respect the symmetry of spacetime, in particular, time translation symmetry and rotational symmetry.
3. Complexity counts the number of gates in tensor network.
4. Contributions to complexity from very small regions are also small.

The most important assumption we use throughout this paper is probably the validity of classical general relativity, at least before very late time.

This paper is organized as follows: In section 2 we give a brief review of some known results about holographic complexity, and then illustrate the idea of symmetry determining complexity by explicit calculations in shockwave geometries. In particular, without a detailed prescription, we'll recover the earlier results about the time dependence of complexity in [1] [2] [4] from the above assumptions. In section 3 we use this idea to show the connection between the time dependence of complexity and transparency of horizons [8]. We first study the connection between the boost symmetry breaking and collision energy in general relativity. Then we revisit the Gedanken experiment by Susskind, Thorlacius in [11] and Hayden, Preskill in [12], and discuss the relations between quantum cloning

²In these four assumptions, assumption 2 essentially implies assumptions 1, 3, and 4. I thank Daniel Harlow for pointing this out.

and firewalls. In section 4 we introduce the two-tape picture, and give an interpretation of firewalls as tape locking. Section 5 points out some unanswered questions.

For the convenience of the reader we'll list some conventions used throughout the paper.

- D refers to the spacetime dimension of the bulk.
- Our convention of time will be that time flows upward on the right side of Penrose diagram, and downward on the left side. See Figure 1. So a precursor $W_L(t_w) = e^{iH_L t_w} W_L e^{-iH_L t_w}$ acting on the left CFT from the remote past would have large positive t_w .
- We'll use boost symmetry and time translation symmetry interchangeably.

2 Time dependence of complexity and boost symmetry

2.1 Known results about holographic complexity

Given a set of simple gates and a simple reference state, the complexity of a state is defined as the minimum number of gates needed to prepare the state. For a continuum field theory, we still don't have a rigorous definition of complexity, but we expect there exists a similar quantity which characterizes how hard it is to prepare a state starting from a particular simple reference state (thermofield double for a two-sided black hole). On the bulk side, it was proposed that complexity is dual to the volume of the maximal surface anchored on the boundary, with some prefactor involving the cosmological constant [1] [2]. Later, it was proposed that complexity equals the action of the Wheeler-DeWitt patch in the bulk [3] [4], and a universal rate of complexity increase was found. Both proposals were tested in various examples.

For example, take the time evolution of the thermofield double:

$$|\psi(t_R, t_L)\rangle = \frac{1}{\sqrt{Z}} \sum_n e^{-\frac{\beta}{2} E_n} |n\rangle_L |n\rangle_R e^{-iE_n(t_R - t_L)}, \quad (2.1)$$

whose holographic dual is a two-sided black hole in Anti-de Sitter space [13] (Figure 1). Since black hole dynamics is chaotic, we don't expect shortcuts (fast forwarding in [14]) before exponentially long time. The complexity will increase linearly with time after an

initial transient period of order the thermal time [1] [4].

$$\mathcal{C}(t_R, t_L) = \mathcal{C}(|\text{TFD}\rangle) + C|t_R - t_L| \quad (2.2)$$

We can also consider more complicated states by perturbing the system at various times. If we act on the state with a precursor $W_L(t_w) = e^{iH_L t_w} W_L e^{-iH_L t_w}$, where W_L is some thermal scale perturbation, it amounts to throwing in a thermal quantum at early time t_w . In the dual geometry there is a shockwave lying close to the horizon [15]. What's the effect of such a perturbation? For a black hole of size l_{AdS} , we expect this quantum takes scrambling time $t_* = \frac{\beta}{2\pi} \log S$ to affect the entire system, during which the perturbation has little effect on complexity. After that the complexity will increase linearly. Bulk calculations indeed give [1] [4]

$$\mathcal{C}(t_R, t_L) = \begin{cases} \mathcal{C}(|\text{TFD}\rangle) + C|t_R - t_L| & t_w - t_R < t_* \text{ or } t_w - t_L < t_* \\ \mathcal{C}(|\text{TFD}\rangle) + C(2(t_w - t_*) - t_R - t_L) & t_w - t_R > t_* \text{ and } t_w - t_L > t_*, \end{cases} \quad (2.3)$$

where in the first line it behaves the same as the unperturbed system. The subtraction of $2t_*$ in the second line is called the switchback effect [1] [16] [17]. The switchback effect is due to cancellations between forward and backward time evolutions during which the perturbation has not yet affected much of the system. Also, as was pointed out in [1] [8], in the range $t_R < t_w - t_*$, $t_L < t_w - t_*$, the complexity always decreases with t_R . In contrast, without any perturbations the complexity will increase with t_R as long as $t_R > t_L$. This is what we mean by abnormal time dependence. We will use this to diagnose transparency of horizons in section 3 [8].

We can also consider insertion of multiple precursors:

$$\begin{aligned} & W_n(t_n) W_{n-1}(t_{n-1}) \dots W_2(t_2) W_1(t_1) \\ &= e^{-iH t_n} W_n e^{iH(t_n - t_{n-1})} W_{n-1} e^{iH(t_{n-1} - t_{n-2})} \dots W_2 e^{iH(t_2 - t_1)} W_1 e^{iH t_1}, \end{aligned}$$

where the perturbations are well separated, $|t_i - t_{i+1}| > 2t_*$, so that the spreadings of different precursors do not interfere. See Figure 2. The arrows indicate the direction of time evolution.

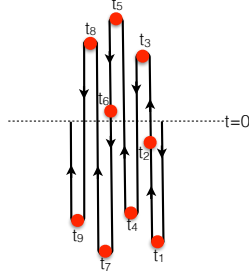


Figure 2: Product of multiple precursors

From considerations of quantum circuits, we expect one switchback from each turn, so the complexity of this operator will be

$$\mathcal{C} = C(t_f - 2n_{sb}t_*), \quad (2.4)$$

where t_f is total folded time, and n_{sb} is the number of switchbacks. On the gravity side the system is dual to black hole geometries perturbed by multiple shockwaves separated by large time [18]. The result (2.4) is reproduced in gravity calculations [1] [4].

In previous examples we considered systems dual to black holes of size l_{AdS} . We can also consider a bigger system with position-dependent perturbations originating at position x_w . Now the perturbation also spreads ballistically. In [19] [20] [2] geometries with localized shocks were studied in detail, and it was shown that in the complexity there will be more complicated position-dependent cancellations [2]. Roughly speaking, the perturbation spreads with butterfly velocity v_B , so the degrees of freedom at position x feel the effect of the perturbation at time $t_w - \frac{|x-x_w|}{v_B}$. This means the contribution to complexity from site x will be proportional to $2 \left(t_w - \frac{|x-x_w|}{v_B} - t_* \right)$.

In all the above calculations, the time dependence of complexity shows certain robust features, like linear increase at large time, switchback delay time of $2t_*$, and abnormal decrease when the interior contains shockwaves. In the rest of this section, we'll rederive these results just from general assumptions 1 to 4. We'll see that, in the bulk, these features are essentially consequences of symmetry and a particular pattern of symmetry breaking which results from the Rindler nature of horizons.

2.2 Linear increase of complexity: Boost symmetry across the horizon

Let's first look at the time evolution of the thermofield double without perturbations (2.1). The state $|\psi(t_R, t_L)\rangle$ describes the Wheeler-DeWitt patch at time (t_R, t_L) , i.e., the bulk

region that is spacelike separated from the chosen boundary time [21]. From assumption 1, complexity is dual to a certain geometric quantity defined within the WDW patch.

Earlier we mentioned boost symmetry across the horizon. Here we'll make this statement precise. In Figure 1, if t_R increases by Δt , the right boundary of WDW patch will move by Δt along any $r = \text{constant}$ curve. This is more easily seen in Eddington-Finkelstein coordinates: Define ingoing time: $dv^* = dt + \frac{dr}{f(r)}$. Then the metric becomes

$$ds^2 = -f(r)dv^{*2} + 2drdv^* + r^2 d\Omega_{D-2}^2 \quad (2.5)$$

These coordinates cover right exterior and future interior regions in the Penrose diagram, and v^* is a good time coordinate on the right future horizon. The boost symmetry in these two quadrants can now be written as

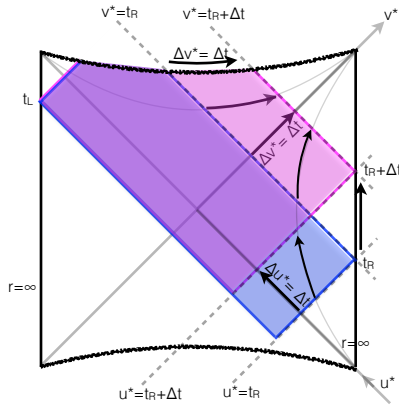
$$v^* \rightarrow v^* + \Delta t. \quad (2.6)$$

We can similarly define $du^* = dt - \frac{dr}{f(r)}$ and it covers the right exterior region as well as the white hole interior.

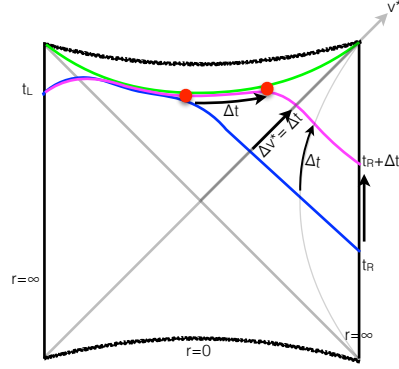
We can intuitively view v^* as a trajectory of a light beam sent from boundary at time t_R along radial directions. The beam will stay on the boundary of the WDW patch $v^* = t_R$. As t_R increases, the spacetime region between $v^* = t_R$ and $v^* = t_R + \Delta t$ is scanned by the beam, and the time dependence of complexity is seen in this scan. As we'll make precise later, the scan can tell us whether the interior is growing uniformly, which shows whether the dynamics of making the state is going well without being disrupted. We should emphasize here, the scanning beam is merely a geometric analogy. All we want is something representing boundary of WDW patch. We don't mean physically throwing in any quanta.

Also note that, even though we work with thermofield double as an example, what we mean by boost symmetry in this paper is a property near the horizon and does not rely on the existence of the other side.

Let's fix t_L and just evolve t_R , when $t_R - t_L$ is large, we see a pattern emerge.



(a) Two WDW patches at different t_R with t_L fixed.



(b) Two maximal surfaces at different t_R with t_L fixed.

Figure 3: Bulk duals of $|\psi(t_R, t_L)\rangle$ and $|\psi(t_R + \Delta t, t_L)\rangle$.

Figure 3a shows WDW patches at two times. As we increase t_R , we move the upper right boundary of WDW patch from $v^* = t_R$ to $v^* = t_R + \Delta t$, and the lower right boundary from $u^* = t_R$ to $u^* = t_R + \Delta t$. The black hole interior is expanding, while the white hole interior is shrinking. With large $t_R - t_L$, the white hole interior is small, and we expect it to have less and less effect on the complexity (assumption 4). The change in the WDW patch is basically to add a piece of Figure 4a to the interior region. Regions like this will be the basic building blocks for our following discussions. They have the feature of being uniform in the Schwarzschild time direction. Look at Figure 4b. We divided it into two parts. The two parts are related by a symmetry transformation $v^* \rightarrow v^* + \Delta t$, so we expect them to give equal contributions to complexity. As a consequence, the contribution of a region like this to complexity will be proportional to Δv^* .³ From now on we will denote it by $C\Delta v^*$:

$$\Delta \mathcal{C} = C\Delta v^* \quad (2.7)$$

This symmetry is the reason for which we find linear increase of complexity. But of course, without a specific prescription we cannot fix the coefficient C .

³In complexity-action prescription, additivity of action is a subtle issue, which was discussed in [22].

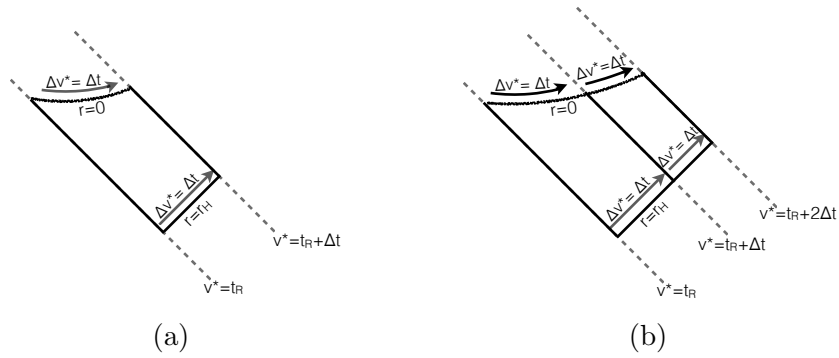


Figure 4: Increase of WDW patch as right time increases from t_R to $t_R + \Delta t$.

The above argument does not depend on the detailed prescriptions of the bulk dual of complexity. It only requires that the prescription respects time translation symmetry, i.e., we don't use different prescriptions at different times! For example, we draw maximal volume surfaces in Figure 3b. The maximal volume surfaces hug the limiting surface (green line), until they separate (red dots). The volume increases linearly with the separation time, and that time is synchronized (up to an additive constant) with boundary time t_R as a consequence of the above mentioned symmetry.

There will be a transient period of order the thermal time, during which the above symmetry argument does not apply. It happens when the tip of the WDW wedge leaves the singularity and the corresponding piece of the interior starts to shrink. We don't expect this period to give big contributions to complexity. Also, we are interested in time scales much bigger than the thermal time, so we will ignore this transient. The transient period ends when the shrinking interior piece becomes very small, or the maximal surface touches the limiting surface. After that the robust linear increase kicks in, and the corrections continue to die off exponentially. Symmetry dictates the linear time dependence in (2.2).

2.3 Breaking boost symmetry: Abnormal time dependence of complexity

What if we perturb the state by inserting a precursor $W_L(t_w) = e^{iHt_w} W_L e^{-iHt_w}$ with t_w large? The geometry was studied in [15]. Now there is a shockwave lying close to the horizon, and it would break the boost symmetry in a specific pattern, see Figure 5.

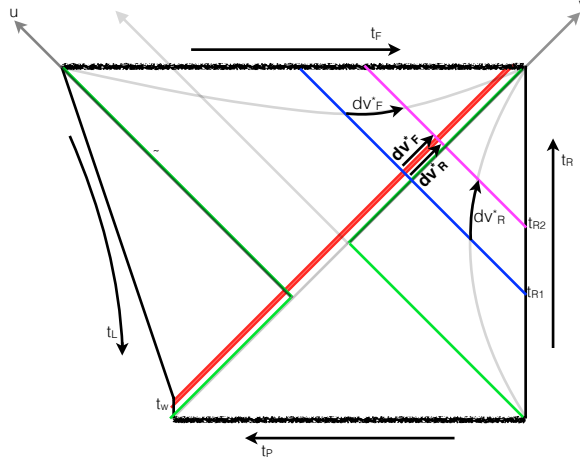


Figure 5: Two sided black hole with one spherically symmetric shockwave

From earlier discussions, we see that what determines the time dependence of complexity is how the interior portion of the WDW patch changes as we vary the boundary time, i.e., as the detecting beam scans the interior it would keep a record of what's going on. Now with a shockwave present, the scanning beam will be pushed forward, by an amount depending on the position where the beam meets the shockwave. Mathematically, ∂_{v^*} is no longer a Killing vector. The time separation Δv^* between the two beams inside horizon will no longer be same as that outside. In Figure 5, consider two ingoing null geodesics from the right boundary with time separation $\Delta t_R = \Delta v_R^*$. Across the shockwave, there will be relative time dilation, and Δv^* will jump when the two geodesics pass the shell.

To quantify this effect, note that near the horizons, the geometry is Rindler-like on both sides of the shockwave:

$$ds^2 = -f(r)dv^{*2} + 2drdv^* + r^2 d\Omega_{D-2}^2$$

where $f(r) = \frac{4\pi}{\beta}(r - r_H)$. Across the shell, there is discontinuity in v^* coordinate:

$$\frac{dv_F^*}{dv_R^*}(t_R) = \frac{f(r_w)}{\tilde{f}(r_w)} = \frac{r_H - r_w}{\tilde{r}_H - r_w} = \frac{r_H - r_w}{\delta r_H + r_H - r_w},$$

where r_w is the radius at which the ingoing null line crosses the shockwave, and δr_H is the increase of horizon radius due to the extra quantum. If we throw in a thermal scale quantum, we have

$$\delta S \sim 1, \quad \frac{\delta r_H}{r_H} = \frac{1}{D-2} \frac{\delta S}{S}.$$

Rindler geometry also gives

$$r_w - r_H = -\frac{\beta}{4\pi} e^{-\frac{2\pi}{\beta}(t_w - t_R - 2R^*)},$$

where $\frac{R^*}{\beta}$ is an order one constant depending on the asymptotic geometry and dimension. So we see the dilation factor is given by

$$\frac{dv_F^*}{dv_R^*}(t_R) = \frac{1}{1 + ce^{\frac{2\pi}{\beta}(t_w - t_R - t_*)}}, \quad (2.8)$$

where $c = \frac{1}{D-2} \frac{4\pi r_H}{\beta} e^{-\frac{4\pi}{\beta} R^*} \delta S$ is an order one constant.

This time dilation factor quantitatively characterizes to what extent the boost symmetry is broken. Note that to arrive at this factor (2.8) we only used the Rindler-like geometry of the horizon. Its functional form is completely robust. Different dimensions or different asymptotic boundary conditions only change the constant c . Let's look at the behavior of this function (Figure 6). When $t_w - t_R < t_*$, $dv_F^* = dv_R^*$, i.e., there is still good symmetry across the horizon. But when $t_w - t_R > t_*$, $dv_F^* = e^{-\frac{2\pi}{\beta}(t_w - t_* - t_R)} dv_R^*$ is exponentially smaller than dv_R^* , i.e., as we increase t_R the future interior does not grow accordingly. We've seen that the linear increase of complexity is a result of the steady expansion of the black hole interior. Now this symmetry is broken, and it gives rise to abnormal time dependence of complexity.⁴ From here we already see that the time dependence of complexity is closely related to the transparency of the horizon, as proposed by Susskind in [8]. We'll explore this more in sections 3 and 4.

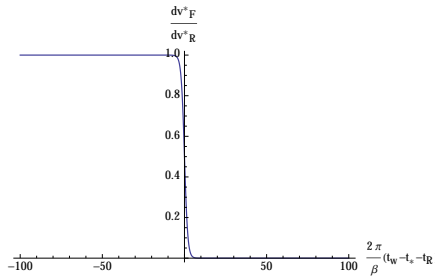


Figure 6: Eddington-Finkelstein time dilation factor as we go across the shockwave, as a function of boundary time.

Because the perturbation at time t_w has very low energy, it causes very little change

⁴In this type of perturbation we considered here (low energy, early time), the symmetry is broken in a particularly simple pattern. The only effect of the shockwave is to give a relative time dilation across the horizon. The change of interior always has the same shape as in Figure 4. The only difference is, Δv^* inside horizon can be smaller than Δv^* outside, which will give rise to a shift to the time dependence of complexity.

to the system right after the perturbation. In order to explore the time dependence of complexity, let's begin with boundary time t_L right above t_w . (In our convention this means t_L is slightly smaller than t_w .) Let $t_R = t_L = t_w$. In Figure 7a, the dual WDW patch is shown in blue.

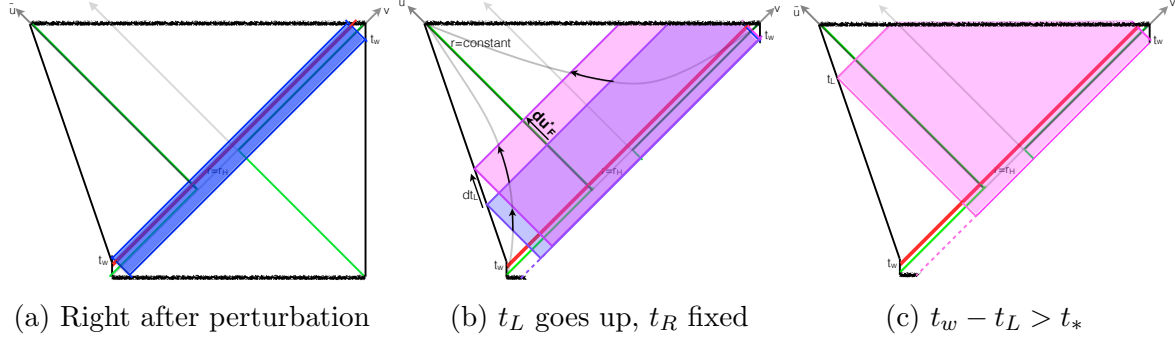


Figure 7: Evolution of WDW patch as we increase left time.

Next, let's push t_L upward with t_R fixed (Figure 7b), i.e., decrease t_L , until $t_w - t_L > t_*$ (Figure 7c). The change in the white hole interior will only show up as some transient behavior for thermal time, so we ignore it. The black hole interior steadily expands. We have

$$\begin{aligned}
 \Delta \mathcal{C}_1 &= C \Delta u_F^* \\
 \frac{d\mathcal{C}_1(t_L, t_R = t_w)}{dt_L} &= C \frac{du_F^*(t_L)}{dt_L} = -C \\
 \mathcal{C}_1(t_L, t_R = t_w) &= \mathcal{C}(|\text{TFD}\rangle) + C(t_w - t_L)
 \end{aligned} \tag{2.9}$$

Thus we see the complexity increases as we push t_L upward.

Now, we fix the left time at t_L , and move t_R downward. See Figure 8.

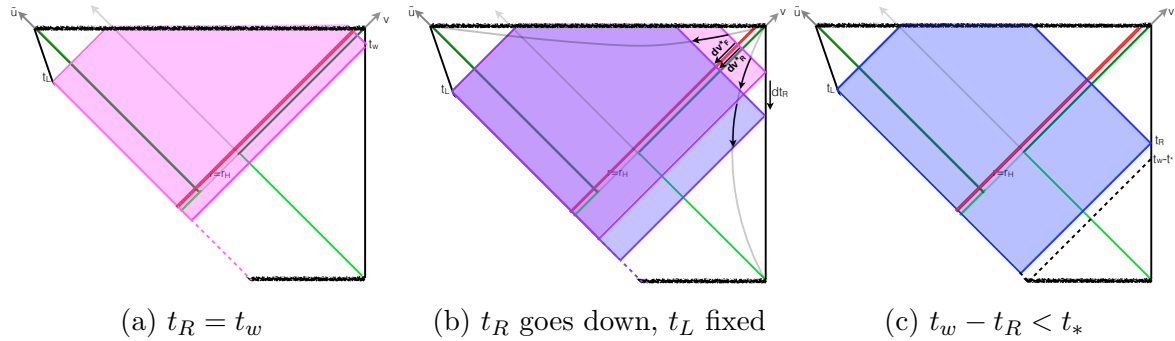


Figure 8: Evolution of WDW patch as we decrease right time.

We start from Figure 8a, where the complexity is given by (2.9). With $t_w - t_R < t_*$ (Figure 8b), we can still ignore the white hole interior. The black hole interior is shrinking, but here, we encounter the shockwave, so we need to take into account the time dilation by (2.8).

$$\begin{aligned}
\Delta \mathcal{C}_1 &= C \Delta v_F^*, \quad dv_R^* = dt_R \\
\frac{d\mathcal{C}_1(t_L, t_R)}{dt_R} &= C \frac{dv_F^*}{dv_R^*}(t_R) = \frac{C}{1 + ce^{\frac{2\pi}{\beta}(t_w - t_* - t_R)}} \\
\mathcal{C}_1(t_L, t_R) &= \mathcal{C}_1(t_L, t_w) + C \left(t_R + \frac{\beta}{2\pi} \log \left(1 + ce^{\frac{2\pi}{\beta}(t_w - t_* - t_R)} \right) \right) - Ct_w \\
&= \mathcal{C}(|\text{TFD}\rangle) + C(t_R - t_L)
\end{aligned} \tag{2.10}$$

In this regime ($t_w - t_R < t_*$, Figure 8c), the symmetry across the horizon is still approximately satisfied, and the complexity increases linearly with the right boundary time. So far we haven't encountered any abnormal behaviors of complexity.

Next, let's push t_R down further, i.e., consider when $t_L < t_w - t_*$ and $t_R < t_w - t_*$. See Figure 9.

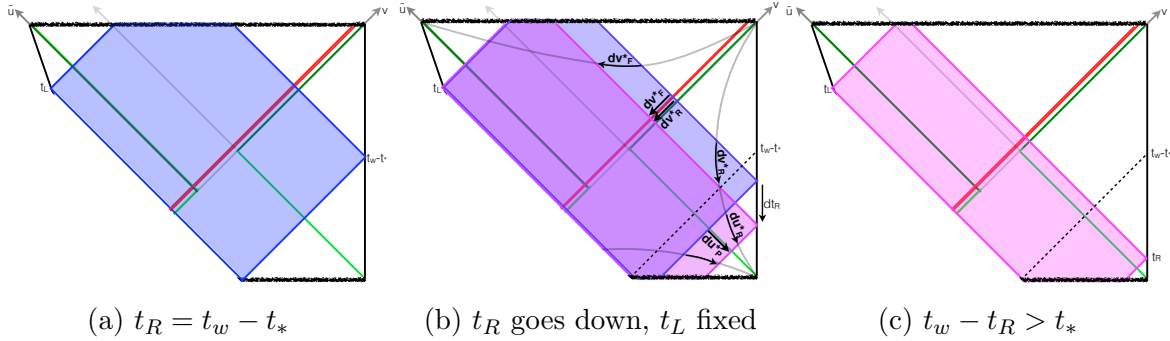


Figure 9: Evolution of WDW patch as we decrease right time.

Now interiors of both the black hole and the white hole change (Figure 9b). We need to consider the contributions to the complexity from both regions. Here, when we talk about the contributions to the complexity from different regions, we are implicitly using

assumption 2 that the tensor network has locality down to AdS scale. We have

$$\begin{aligned}
\Delta\mathcal{C}_1 &= C\Delta v_F^* - C\Delta u_P^* \\
\frac{d\mathcal{C}_1(t_L, t_R)}{dt_R} &= C \left(\frac{dv_F^*}{dv_R^*}(t_R) - \frac{du_P^*}{dt_R} \right) = -C \frac{ce^{\frac{2\pi}{\beta}(t_w - t_* - t_R)}}{1 + ce^{\frac{2\pi}{\beta}(t_w - t_* - t_R)}} \\
\mathcal{C}_1(t_L, t_R) &= \mathcal{C}_1(t_L, t_w - t_*) + C \frac{\beta}{2\pi} \log \left(1 + ce^{\frac{2\pi}{\beta}(t_w - t_* - t_R)} \right) \\
&\quad - C \frac{\beta}{2\pi} \log \left(1 + ce^{\frac{2\pi}{\beta}(t_w - t_* - (t_w - t_*))} \right) \\
&= \mathcal{C}(|\text{TFD}\rangle) + C(2(t_w - t_*) - t_L - t_R)
\end{aligned} \tag{2.11}$$

Without perturbations, the answer would be $\mathcal{C}(|\text{TFD}\rangle) + C|t_R - t_L|$. We see that the time dependence becomes quite different in this regime. For example, let's set $t_L = -\infty$ ⁵. Without perturbations, the complexity will always increase with t_R . Now with a shockwave present, the complexity will decrease with t_R when $t_w - t_R > t_*$. This abnormal time dependence is a consequence of breaking of boost symmetry as we go across the shockwave.

Combining (2.9), (2.10), and (2.11), we recover the time dependence in (2.3).

In $D > 3$, because the asymptotic boundaries bend outward, there will be time of thermal length when the complexity is not changing. We don't understand the reason. In this paper we'll consider it as part of the transient behavior and ignore it in our calculations.

In the above calculations, we ignored contributions from transient periods. It will depend on specific geometries and cannot be completely determined by symmetry argument. However, in calculating the complexity of the precursor $W(t_w)$, the transient behavior has little effect. We carry out the calculation in Appendix A.1, and see a smooth exponential-to-linear transition as one expects from analysis of quantum circuits.

2.4 More examples

The idea that boost symmetry determines the time dependence of complexity should be clear from the above calculations. Following the same procedure we can consider more complicated examples. For states perturbed by multiple precursors, we follow the time evolution of preparing the state. Each time we make a perturbation, it takes a scrambling time for the memory of the previous perturbation to disappear, and the new perturbation

⁵This is just a technical choice to avoid transient period. We'll give more explanations about this in section 3.

takes a scrambling time to be squeezed up against the horizon. This is the origin of the $2t_*n_f$ subtractions in (2.4). The detailed calculations are in Appendix A.2.

For states with localized perturbations, the time dilation factors will be position dependent. From a symmetry point of view, not only the boost symmetry is broken, but the translation symmetry in the spatial transverse direction is also broken. But again, at a coarse-grained level it's broken in a particularly simple pattern. More details are in Appendix A.3. When we vary the transverse positions of the scanning beam⁶, we will also detect the transverse variations of geometry caused by the propagation of perturbations. Intuitively, one can imagine a tensor network lying at the stretched horizon. It evolves as the dynamics goes on. When we look at the complexity history, we scan this network. In particular, from the profile of complexity we can see some part of the network is normal while some part is disrupted by high energy. This will be the topic of the next section: diagnosing the horizons by complexity.

3 Time dependence of complexity as a diagnosis of good interior

From earlier discussions it should be already clear why we can use the complexity to detect the interiors of black holes or white holes. When there is nothing abnormal, the interior steadily expands as the boundary time evolves, and this will give linear increase of complexity. An abnormal decrease of complexity means this evolution is disrupted.

There is one technical point here. There are four horizons in the example we discussed, right / left future and past horizons. If we want to detect the transparency of one particular horizon, we don't want to be in transient period, i.e., we want the boundary of WDW patch to touch the singularity. To achieve this, we send the other time to infinity. For example, if we want to detect the transparency of the right future horizon, we set $t_L = -\infty$, i.e., push the left time up to the upper corner. See Figure 10. As we increase the right time, a linear increase in complexity means the part just scanned has a good interior (Figure 10a). If the complexity decreases somewhere, it means there is something wrong in that part of the interior (Figure 10b).

⁶For localized perturbations, scanning light beam is not a perfect analogy. With localized shockwaves, null geodesics do not stay on constant spatial positions. Nevertheless we want to detect the interior along slices with constant spatial positions in order to compare with the case without perturbations.

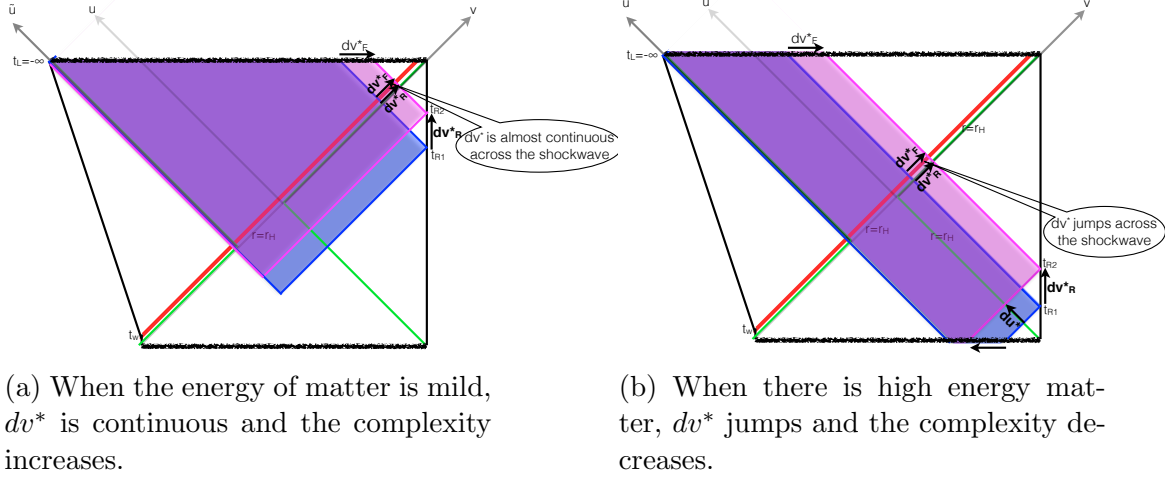


Figure 10: Diagnose the smoothness of right future horizon. Set t_L at the top and scan the interior by increasing Eddington-Finkelstein coordinates $v^* = \text{constant}$.

If we want to detect the transparency of the right past horizon, we instead put the left time at the lower corner ($t_L = \infty$), and scan the interior with Eddington-Finkelstein time $u^* = \text{constant}$, see Figure 11.

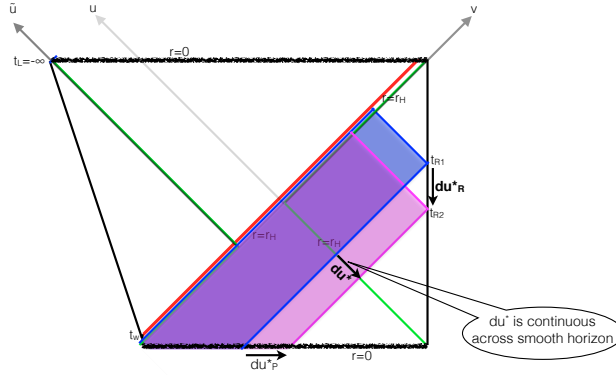


Figure 11: Diagnose the smoothness of right past horizon by setting t_L at the bottom and scan the interior by Eddington-Finkelstein coordinates $u^* = \text{constant}$.

Similarly, one can detect the smoothness of the left horizons by looking at the dependence of complexity on the left time. If we want to detect the upper left horizon, we set $t_R = \infty$. If we want to look at the lower left horizon, we set $t_R = -\infty$.

In the above examples of two-sided black holes, we anchored time of the other side at infinity to avoid transient period. But for black holes with wide enough Penrose diagrams, this is unnecessary (Figure 27, 28). It's certainly unnecessary for one-sided black holes. In those cases, if we move time up, we are detecting the transparency of the upper horizon, while if we move time down, we are detecting the transparency of the lower horizon. Then

a natural conjecture is, for a state with both future and past horizons, at a fixed time you cannot have firewalls on both of them. This is true in the examples we considered here. It's also natural from complexity considerations. The interiors store the quantum circuit making the state. Firewalls lock them up. But the complexity always changes before it reaches the maximum, so there must always be some horizon free of firewalls at work. More discussions on quantum circuit and firewalls are in section 4.3.

3.1 Boost symmetry breaking and firewalls

We've seen that the time dilation factors quantitatively characterize to what extent the boost symmetry is broken. They control the time dependence of complexity. In this subsection we'll show they are also directly related to the collision energy experienced by an infalling thermal quantum and diagnose firewalls.

Assume there is a homogeneous perturbation, as in section 2.3, Figure 5. Let's say we throw in another thermal scale quantum from right boundary at time t_R . It meets the matter shell at radius r_w . We have

$$\frac{dv_F^*}{dv_R^*}(r_w) = \frac{r_H - r_w}{\delta r_H + r_H - r_w} = \frac{(r_H - r_w)/\delta r_H}{1 + (r_H - r_w)/\delta r_H}$$

where for a thermal scale quantum, $\delta r_H \sim \frac{r_H}{S} \sim \frac{l_p}{S^{(D-3)/(D-2)}}$ which is sub-Planck scale. On the other hand, the collision energy will be

$$\begin{aligned} \frac{E_c^2}{m_p^2} &= \frac{4E_1 E_2}{m_p^2 f(r_w)} \sim \frac{T l_p^2}{r_H - r_w} \sim \frac{T S l_p^2}{r_H} \left[\left(\frac{dv_F^*}{dv_R^*} \right)^{-1} - 1 \right] \\ &\sim \frac{r_H}{\beta} S^{\frac{D-4}{D-2}} \left[\left(\frac{dv_F^*}{dv_R^*} \right)^{-1} - 1 \right] \propto \left(\frac{dv_F^*}{dv_R^*} \right)^{-1} - 1 \end{aligned} \quad (3.12)$$

We see that with significant time dilation: $\frac{dv_F^*}{dv_R^*} \ll 1$, the collision-energy-squared is inversely proportional to the dilation factor and can reach the Planck scale.

More generally, in appendix B.1 we use Raychaudhuri equation to relate the time dilation factors to the stress-energy tensors. Stress-energy tensors cause geodesics to focus. This fact can be translated into time dilation factors which control the time dependence of complexity.

There is one subtlety here. What do we mean by firewalls? Stress-energy tensors surely signal firewalls. They characterize energy from matter. But gravitons also carry energy.

In fact, in presence of localized perturbations, there can be nontrivial time dilation factors at places where the energy momentum tensor $T_{\mu\nu}$ vanishes. Someone who tries to cross the horizon there will still be hit hard, not by matter, but by gravitational shockwaves. We should also consider them as firewalls even though the local stress-energy tensors vanish. To quantitatively describe this effect, we can use the Landau-Lifshitz pseudotensor [23].

The Landau-Lifshitz pseudotensor characterizes the energy-momentum flows carried by gravitons. It is only a tensor under linear coordinate transformations, in particular, Lorentz transformations. To describe the experience of an infalling observer, we should evaluate it in the local inertial frame right before the observer crosses the horizon. In appendix B.2, we show that in the geometries we considered in this paper, the collision energy is always controlled by the time dilation factors as in (3.12), no matter it's from collisions with matters or with gravitons.

We see from classical general relativity that the breaking of boost symmetry is a diagnosis of firewalls. Next, we'll take another point of view, and see that these two concepts are closely related within the context of quantum cloning.

3.2 Boost symmetry breaking and quantum cloning: Firewall is necessary

We've seen that a necessary condition of complexity increase is good boost symmetry across the horizon. In this subsection let's explore the operational meaning of this symmetry.

Consider two infalling observers. A good boost symmetry across horizon would guarantee that they have enough separation inside the black hole. Mathematically, their Eddington-Finkelstein time separation stays the same before and after they jump in. This plays an important role in complementarity of black hole description [24]. Consider an infalling observer carrying a spin. Black hole complementarity states that, information about the spin is carried both by Hawking radiations as well as the infalling observer. So information appears both inside as well as outside the horizon, but no one can see the duplication.

Let's revisit the Gedanken experiment by Susskind and Thorlacius in [11]. See Figure 12.

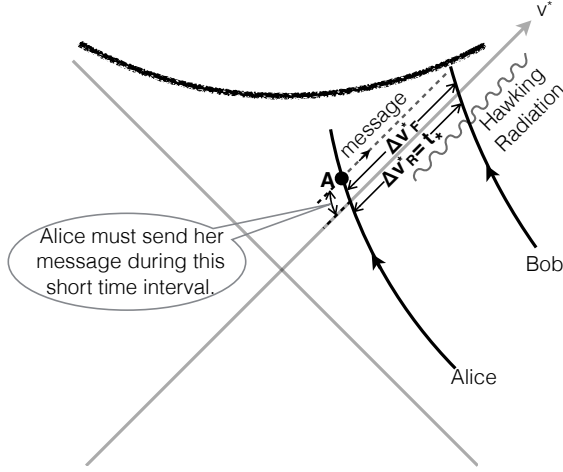


Figure 12: Bob falls into the black hole after collecting enough Hawking radiation to learn about the spin.

Alice falls into the black hole carrying a spin. Bob collects the Hawking radiation, and then also jumps in.⁷ If he encounters the infalling spin, he would see cloning of quantum information, which is a violation of quantum mechanics. We assume Alice falls in after Page time. Hayden and Preskill argued that Bob only needs to wait for scrambling time t_* before the information about the spin comes out [12]. So we assume $\Delta v_R^* = t_*$ where Δv_R^* is the Eddington-Finkelstein time separation between Bob and Alice when they are both outside the horizon. After Bob jumps in, Alice can send him a message about the spin she carries. Alice's message needs to reach Bob before Bob hits the singularity, so the proper time left for Alice to send the message depends on their separation (Figure 12). If we trace the message back, we see that it has to be sent out at a radius r_A with $r_H - r_A \sim r_H e^{-\frac{2\pi}{\beta} \Delta v_F^*}$, where Δv_F^* is their Eddington-Finkelstein time separation inside the horizon. The message needs to be sent within Alice's proper time $\Delta\tau \sim r_H - r_A = r_H e^{-\frac{2\pi}{\beta} \Delta v_F^*}$. By uncertainty principle, the message needs to have energy

$$E_{\text{Alice}} \sim \frac{1}{\Delta\tau} \sim \frac{1}{r_H} e^{\frac{2\pi}{\beta} \Delta v_F^*}$$

To protect Bob from seeing quantum cloning, we want E_{Alice} to be bigger than the

⁷Here we ignore the fact that Bob may need long time to decode Hawking radiation. See [25].

Planck energy m_p , i.e.,

$$\Delta v_F^* > \frac{\beta}{2\pi} \log \frac{r_H}{l_p} = \frac{1}{D-2} t_* = \frac{1}{D-2} \Delta v_R^*. \quad (3.13)$$

A good symmetry across the horizon would guarantee that Δv_F^* is roughly the same as Δv_R^* , so the above condition for no-cloning is satisfied. However, when that symmetry is badly broken, Δv_F^* can be exponentially smaller than Δv_R^* , and condition (3.13) is no longer satisfied. If Bob crosses horizon at this time, Alice is right next to him, and he might be able to see quantum cloning. To say in another way, it's not allowed to enter a non-expanding interior. So we see the necessity of firewalls: when the boost symmetry is broken and the complexity decreases, a firewall is necessary to protect someone from seeing quantum cloning.

In summary, Quantum cloning happens whenever something enters the horizon. Either the interior keeps stretching so no one can see the cloning, or there must be a firewall which prevents the entire experiment from taking place. On the other hand, the interior stores complexity. A constant stretching of the interior means a steady increase of complexity.

4 Quantum circuit and firewalls

4.1 Future and past tapes

Tensor networks support spacetime. As pointed out by Hartman and Maldacena [7], we can consider a tensor network laid near the horizon. As time grows, more layers are added to the tensor network, recording the action of the Hamiltonian. It was pointed out by Susskind that the minimal quantum circuit preparing a state is stored in a tape behind the future horizon (Figure 13) [9]. It's called a tape because it keeps a record of the past actions of making the state. This picture is also the basic reason the complexity-volume duality works. For more details, see section 3.3 and Figure 10, 17, 19 in [9].

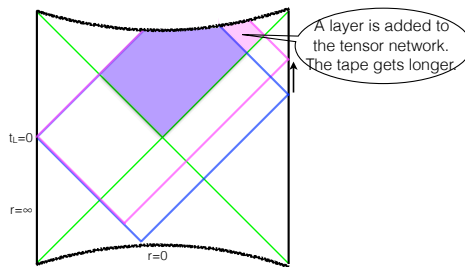


Figure 13: The growth of the future tape

In this section, we'll see that, the future and past interiors are like two tapes storing different parts of the minimal circuit preparing the state⁸. In the following, we'll sometimes call the future interior the future tape, and the past interior the past tape.

Here is a simplest example. Consider the time evolution of thermofield double. We fix $t_L = 0$, and draw the WDW patches as well as the maximal surfaces at different t_R (Figure 14). When $t_R < 0$, the past interior grows to the past (Figure 14a). This is a white hole state, and the minimal circuit preparing this state is stored in the past tape. When $t_R > 0$, the future interior expands to the future (Figure 14b). This is a black hole state, and the minimal circuit is stored in the future tape.

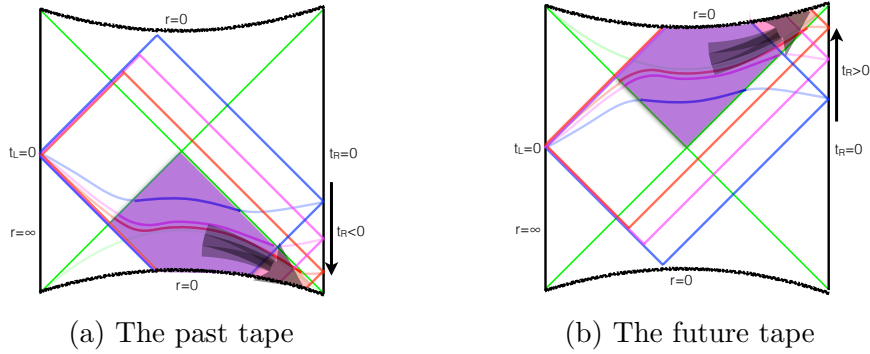


Figure 14

Where are the tapes? There is only AdS distance in the radial direction of the interior. When we talk about tapes, we mean this AdS thickness thing (colored regions in Figure 14), which gets longer along spacelike Schwarzschild time direction. We don't consider the tensor network to be on any specific slice. In complexity-volume duality [1] [2], the maximal volume surface is used as a gauge-invariant way to represent the wormhole. As in Figure 14, when the future/past interior expands, the maximal surface also goes through the future/past interior. But it does not mean that particular slice is physically special. In fact, we can't localize the tensor network to distances smaller than AdS scales.

⁸What we mean by future and past interiors should be clear from the context in the following discussions. Mathematically, we can consider the part of spacetime inside the future apparent horizon as the future interior, while the part inside the past apparent horizon as the past interior.

4.2 Two tape picture: Why are there both future and past horizons?

Let's make a more detailed comparison between the picture of the quantum circuit preparing a state and the dual black hole geometry. We'll see that the minimal circuit preparing a state can contain both forward and backward Hamiltonian evolutions (for example, left and right legs in Figure 16b for a precursor). The forward Hamiltonian evolution part is stored in the future tape, while the backward Hamiltonian evolution part is stored in the past tape.

To avoid complications from the other side, in this section we'll use a one-sided black hole as an example. We start from an AdS black hole formed from a collapsing shell at $t = 0$:

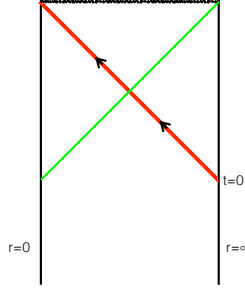


Figure 15: A one-sided black hole formed by a collapsing shell.

Now we perturb the state by inserting a precursor: $W(t_w) = e^{iHt_w} W e^{-iHt_w}$ with $t_w > t_*$. Figure 16a illustrates such a precursor composed of three parts: forward time evolution e^{-iHt_w} , perturbation W , and backward time evolution e^{iHt_w} . There will be cancellations right after the perturbation, and in Figure 16b the blue line represents the minimal circuit after the cancellations. See [1] for more explanations.

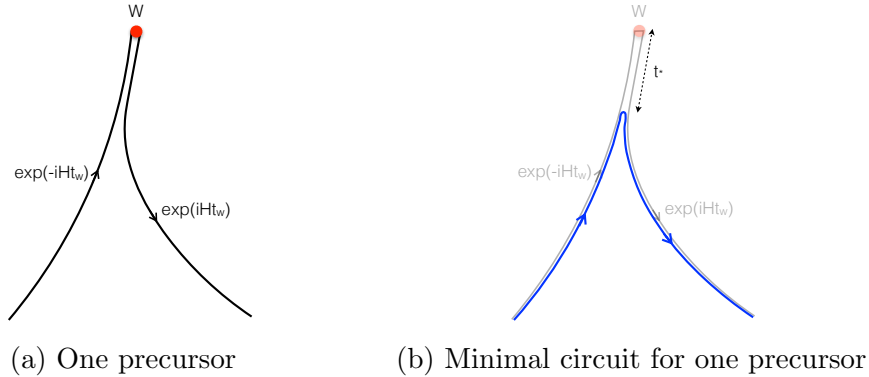


Figure 16

With the perturbation, a past singularity will form, and the dual geometry looks like

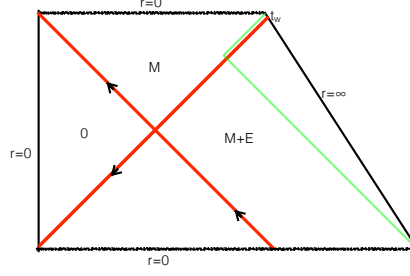


Figure 17: A one-sided black hole perturbed by a precursor

To match the quantum circuit picture with the black hole geometry, we start at $t = 0$ when the black hole just formed (Figure 18a), and time evolve to $t = t_w$ (Figure 18b). The minimal circuit of this forward time evolution is shown as A in Figure 18c. It is stored in the future interior (A in Figure 18b). In this regime, the white hole has not formed yet, and the future tape is working.

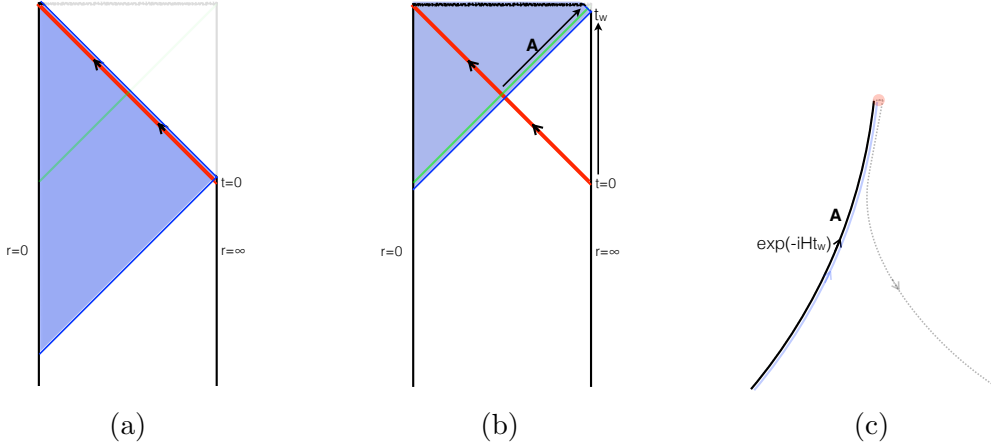


Figure 18

Next, we perturb the state by W at $t = t_w > t_*$ and then evolve backward. We start right after the perturbation at $t = t_w$ (Figure 19a), where the minimal circuit A (Figure 19b) is stored in the future interior (A in Figure 19a). A white hole forms.

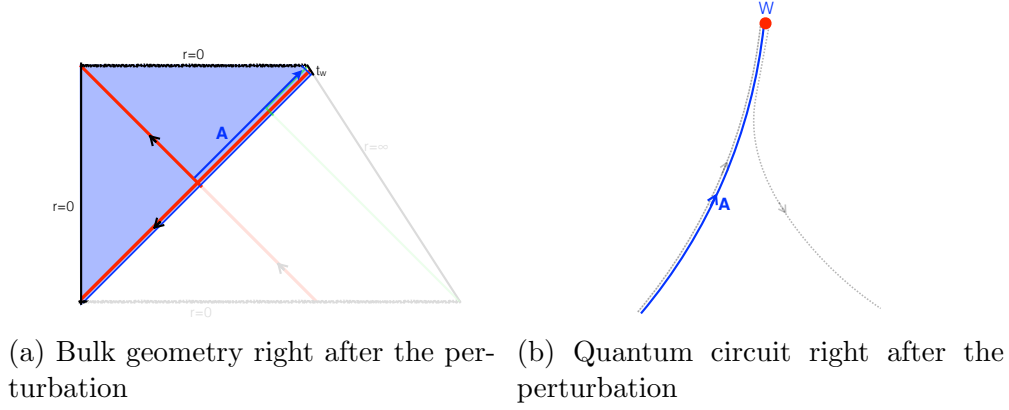


Figure 19

We evolve backward, until $t = t_w - t_*$. When $t_w - t < t_*$, the backward time evolution will cancel part of earlier forward time evolution (In Figure 20b, the circuit in B cancels part of circuit in A, resulting in minimal circuit C) [1]. As a result, the minimal circuit gets smaller, and still only contains forward Hamiltonian evolution (C in Figure 20b). On the bulk side, the future interior shrinks (Figure 20a). We ignore the change in the past interior, since the volume inside the post collision region is small. So we see that the minimal circuit making the state at this time ($t = t_w - t_*$), which contains only forward Hamiltonian evolution, is still stored in the future interior.

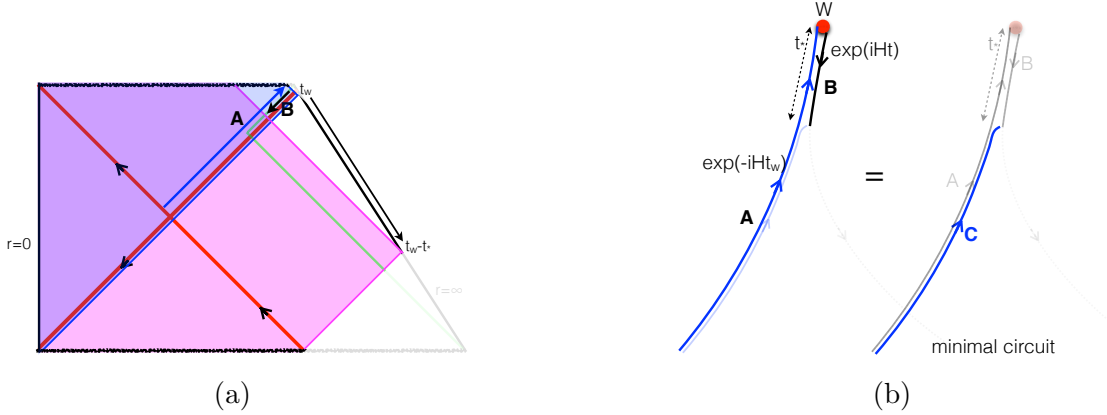


Figure 20

Next, we further decrease time t . By now ($t < t_w - t_*$) the perturbation has spread to the entire system and there are no more cancellations between the backward and forward Hamiltonian evolutions. The minimal circuit is growing from backward Hamiltonian evolution (D in Figure 21b). On the bulk side, the future interior does not change because there is significant time dilation across the firewall (Figure 21a). The past interior

is expanding, and the circuit from backward Hamiltonian evolution is stored there (D in Figure 21a).

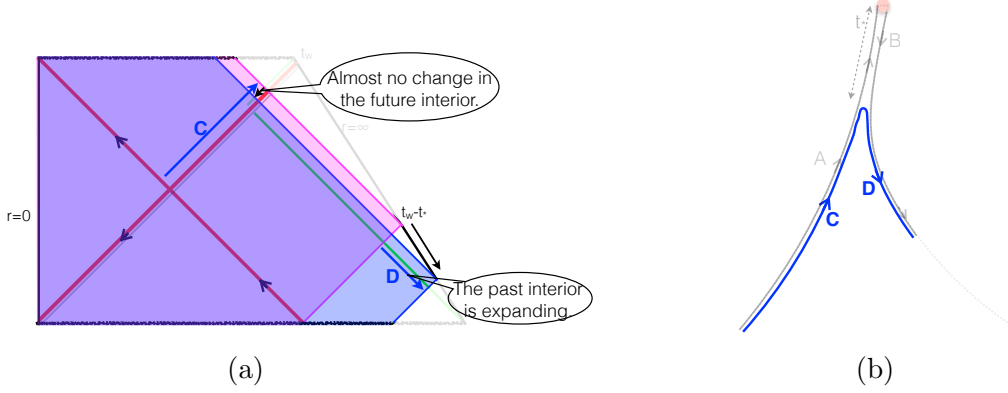


Figure 21

Eventually, we get to $t = 0$, and we have made the state perturbed by the precursor $W(t_w)$. In the minimal circuit making this state, the forward Hamiltonian evolution (C in Figure 22b) is stored in the future interior (C in Figure 22a), while the backward Hamiltonian evolution (D in Figure 22b) is stored in the past interior (D in Figure 22a).

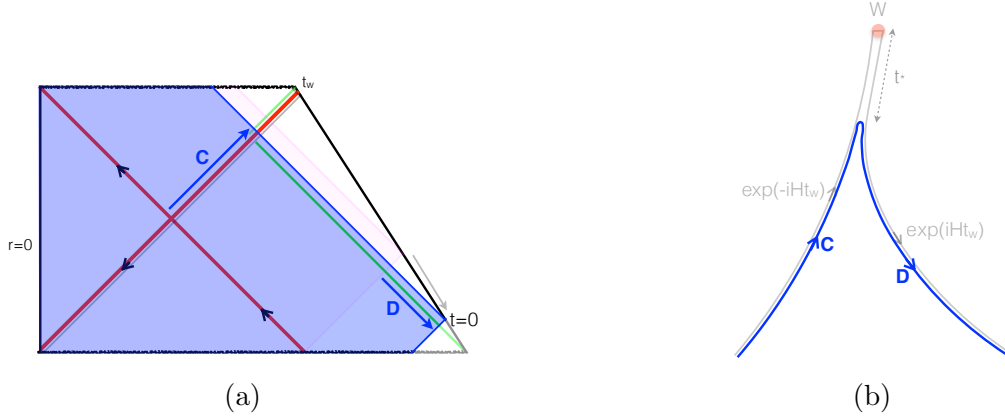


Figure 22

We see the roles played by future and past horizons. While the stretching of the future interior is responsible for the increase of complexity to the future, to have complexity decrease it seems necessary to have past horizons. Complexity is stored in interiors. The existence of past horizons means the complexity can also increase to the past, and so is a manifestation of time reversibility.

4.3 Tape locking: an interpretation of firewalls

There is one confusing point about the relation between the time dependence of complexity and firewalls. The time dependence of complexity is not a conventional property of states in general. There are properties of states described by linear operators, like position, momentum, e.t.c.. Entanglement is also a property of states, though may not always be linear [26]. These properties are independent of dynamics. But given a state, without knowing the Hamiltonian, in general one cannot talk about the time dependence of its complexity. If the existence of firewalls is a state property, then how do we use the time dependence of complexity to detect firewalls? The two-tape picture gives an answer to this question. We'll see that on the quantum circuit side, the existence of firewalls means a particular tape is locked, and we use the time dependence of complexity to detect the locking of tapes.

Let's look at the perturbed one-sided black hole at a particular time $t > t_w - t_*$ (Figure 23a). We've seen that its complexity increases with time. The perturbation is not too close to the horizon, and there is no firewall.

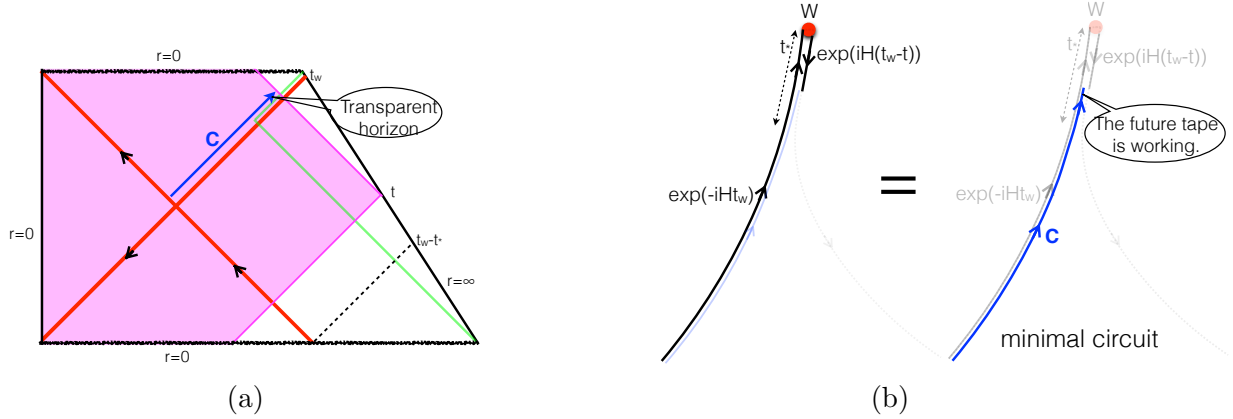


Figure 23

The minimal circuit preparing this state is shown as blue line C in Figure 23b. At the end of the minimal circuit it is forward Hamiltonian evolution. The quantum gates making the state are being laid on the future tape, so we say the future tape is working. This is why the complexity will increase as we increase the time. Because tensor network supports spacetime, this also implies the future interior is growing, and quantum cloning can be protected by this fast stretching. The future horizon is transparent here. The past tape is not working, and inside the past interior there is high energy collision.

If we look at the state at another time $t < t_w - t_*$, the minimal circuit will look completely different:

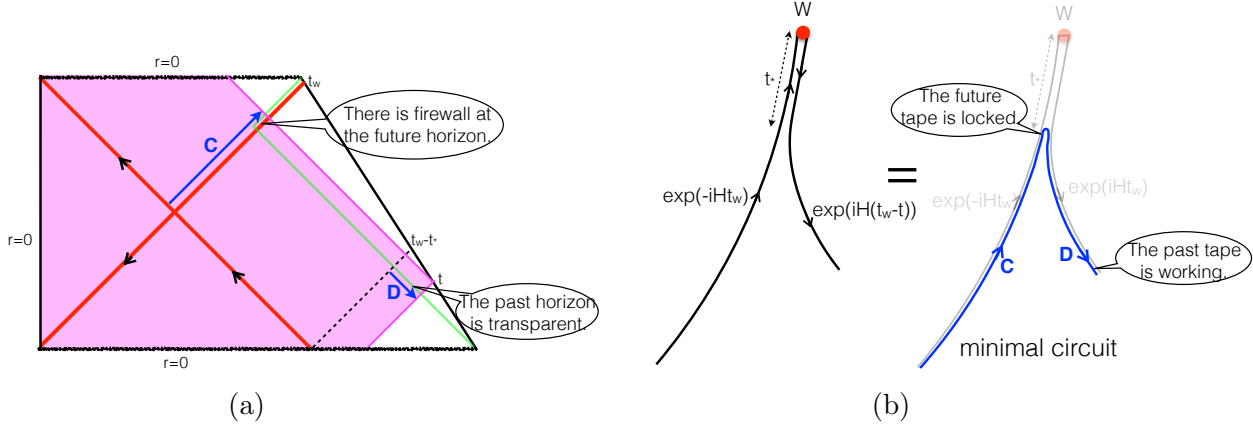


Figure 24

Now the minimal circuit (blue line in Figure 24b) contains both forward Hamiltonian evolution (C in Figure 24b) and backward Hamiltonian evolution (D in Figure 24b). The forward evolution part is stored in the future tape (C in Figure 24a), while the backward evolution part is stored in the past tape (D in Figure 24a). At the end of the minimal circuit it is backward Hamiltonian evolution, i.e., it is the past tape working. No quantum gates are being laid on the future tape at this moment, and we say the future tape is locked. Since the future interior does not grow, if things can still enter the interior at this point, there would not be enough spatial separations between an infalling object and an infalling observer with Hawking radiations in his hands, and the observer could readily see quantum cloning. However, firewalls will save the day. With a firewall present, nothing can enter the interior at this moment. The existence of firewalls prevents cloning from happening. From the complexity point of view, we see the complexity decrease because the past tape working means the complexity is increasing towards the past.

We see that firewalls have a simple interpretation in this picture. The future and past interiors are like two tapes storing different parts of the minimal circuit. Each tape can be in two states: working, when there are quantum gates being laid on it, or locked, when no more gates are being laid on it. When a tape is working, it's getting longer from those newly added quantum gates. So the corresponding interior steadily grows with time, and quantum cloning, if happens, is protected from being seen. At this moment the interior is open for someone to enter. When a tape is locked, no change happens at the end of it. As a result the corresponding interior ceases to grow. Without the fast stretching protecting

an infalling observer from seeing quantum cloning, a firewall will instead ensure nothing can fall in at this moment.

5 Discussion

We showed that the time dependence of complexity is closely related to the boost symmetry across horizons. The Rindler-like nature of horizons guarantees that the boost symmetry is broken in simple and universal forms, which gives rise to the various robust features of the time dependence of complexity.

From both the perspectives of collision energy experienced by an infalling object and quantum cloning, we see that the breaking of this boost symmetry is closely related to firewalls. This in turn shows the connection between the time dependence of complexity and the transparency of horizons [8].

Identifying the black / white hole interiors as tapes storing different parts of the minimal circuit preparing the state, we see the various ideas fit together: the tape working, means the interior growing, the boost symmetry staying good, quantum cloning being well protected, and transparent horizon. On the other hand, a tape being locked, means the interior stops growing, the boost symmetry being broken, no protection for quantum cloning, and the existence of a firewall as a new protection.

It would be interesting to understand this boost symmetry and symmetry breaking in the boundary field theory. Intuitively, without perturbations, there is uniform dynamics. Right after the perturbation, the perturbation is distinguishable and it's not very symmetric. After it is scrambled, locally everywhere is the same, and certain symmetry is regained. A steady increase of complexity is a kind of time translation symmetry.

In our analysis we ignored various regions of AdS size and considered them as giving rise to transient behaviors. However, they serve as the “seeds” for later steady expansion of the interior. On the quantum circuit side, they correspond to the most interesting behavior of switching from one tape to another. But without an understanding of sub-AdS locality, we don't have tools to address them.

Another problem is about linearity of firewalls. If we want to see whether the superposition of two states has a transparent horizon, we need to know something about the minimal circuit preparing the superposition. Unfortunately so far we know very little about the complexity of superpositions without using ancilla qubits.⁹ Maybe a more basic

⁹We thank Scott Aaronson for explanations on this.

question to ask is, on what states can we talk about superpositions? It may not make sense to talk about superpositions of states with large relative complexity. One possibility is that the existence of firewalls behaves like an ordinary property only if we restrict to certain subspace of states, like entanglement entropy as shown by Harlow recently in [26]. Another possibility is that to describe black hole interiors we indeed need some violations of quantum mechanics, but no observer should be able to see it. After all, testing if there exists a firewall or not is not some experiment that can be repeated by a single observer.

There are still many questions surrounding the two-tape picture. It works well in the examples we studied in this paper, but how general is it? How to lock up a tape in general? Can you have firewalls on the future horizon without a white hole in the past? In particular, how should we understand the firewalls Alice would create if she has a powerful quantum computer and more than half of the Hawking radiations [21]? One interpretation is, the large complexity needed for Alice to destroy a transparent horizon [25] plays exactly the role of the complexity stored in the past tape as if there were a white hole in the past. In this case, the past tape is the actual tape in Alice's quantum computer. We'll come back to this two-tape and tape locking picture in future.

A key assumption we used throughout this paper is the validity of classical general relativity, at least until very late time. This assumption might break down earlier. But it's interesting to see how far we can go with these assumptions. We see that all these ideas: complexity, minimal circuit, quantum cloning, and firewalls fit together, at least in the examples we studied. However, we do keep in mind that these examples can be special. Furthermore, it may be a very strong constraint to require that a state determines the minimal circuit making it. To further address this, a better understanding of complexity and minimal circuit is needed.

Acknowledgements

I am grateful to my advisor Lenny Susskind for tremendous help in various aspects. I thank Daniel Harlow, Douglas Stanford for discussions. I thank Scott Aaronson for explanations on complexity of superposition.

A Complexity of perturbed thermofield double

A.1 Complexity of a precursor

Our goal is to calculate the complexity of the precursor $W(t_w) = e^{iHt_w} W e^{-iHt_w}$. Consider the complexity of $W_L(t_w)|\text{TFD}\rangle$. Geometrically, we start from $t_R = t_L = t_w$ (Figure 25a), decrease $t_R = t_L = t$ (Figure 25b) until $t = 0$ (Figure 25c). During this process the tip of WDW patch never leaves the singularity, so the symmetry argument always works.

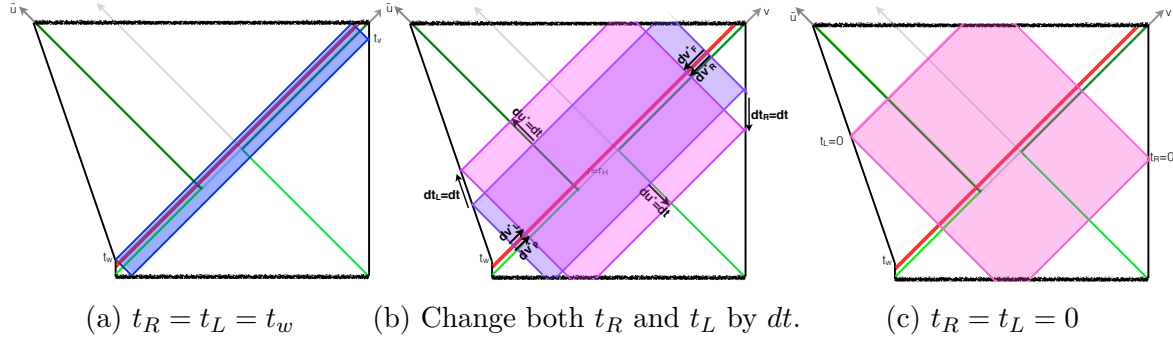


Figure 25: Complexity of a precursor

We have

$$\begin{aligned} \frac{d\mathcal{C}_1(t, t)}{dt} &= C \left(\frac{dv_F^*}{dv_R^*} - 1 + \frac{dv_P^*}{dv_L^*} - 1 \right) = -2C \frac{ce^{\frac{2\pi}{\beta}(t_w - t_* - t)}}{1 + ce^{\frac{2\pi}{\beta}(t_w - t_* - t)}} \\ \mathcal{C}_1(t, t) &= \mathcal{C}(|\text{TFD}\rangle) + 2C \frac{\beta}{2\pi} \log \left(1 + ce^{\frac{2\pi}{\beta}(t_w - t_* - t)} \right) \end{aligned}$$

With $t = 0$, we get the complexity of a precursor:

$$\mathcal{C}(W(t_w)) = 2C \frac{\beta}{2\pi} \log \left(1 + ce^{\frac{2\pi}{\beta}(t_w - t_*)} \right). \quad (\text{A.14})$$

This has exactly the same functional form as geodesic distance in the BTZ black hole [15]. The symmetry argument will be accurate when t_w is a fraction of scrambling time (when Rindler approximation is good). Since this calculation only depends on the boost symmetry and the Rindler nature of horizons, it must reflect some universal property of horizon dynamics. It's interesting to see the exact functional form also come out of a simple epidemic model [16].

A.2 Switchback from multiple precursors

In section 2.3 we studied in detail the case of one precursor, and saw that the time dependence of complexity is determined by the behavior of boost symmetry across the horizon. Let's see how this works with multiple perturbations separated by large time: $W(t_{w2})W(t_{w1})|\text{TFD}\rangle$. We assume $t_{w1} > t_*$, and $t_{w2} < -t_*$. See Figure 26. [18]

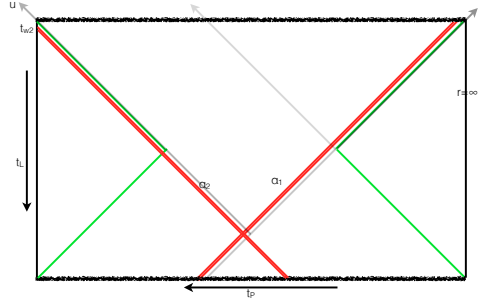


Figure 26: Thermofield double perturbed by two precursors separated by large time

At $t_L = t_{w2}$, the effect of the second perturbation is small and we essentially have a state with one shockwave (Figure 27a). We have its complexity from previous section (2.9), (2.10), and (2.11).

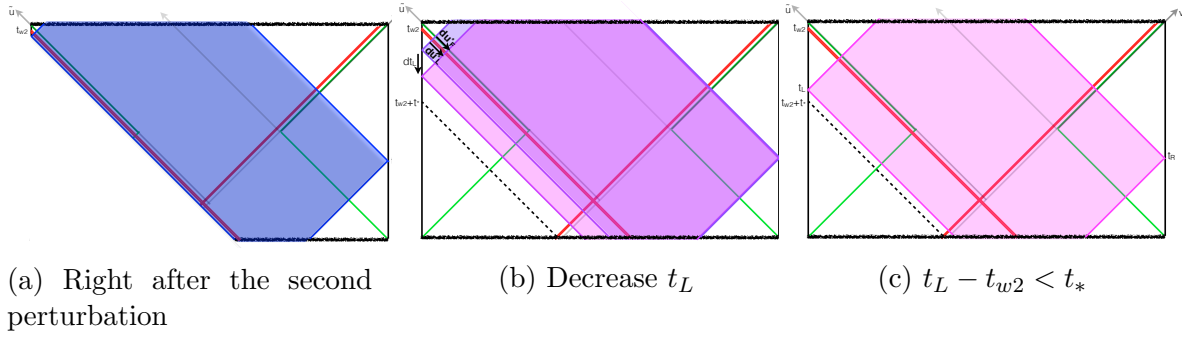


Figure 27: Less than scrambling time after the second perturbation

Now we evolve t_L downward. The black hole interior shrinks, but inside the white hole, we'll encounter a post collision region (Figure 27b). The collision energy increases exponentially with $t_{w1} - t_{w2}$. With $t_{w1} - t_{w2} > 2t_*$, it will be comparable to the mass of the black hole, but the size of spacetime region there is exponentially small. Here, we use assumption 4 to ignore its contribution.¹⁰ It takes scrambling time for the first shockwave

¹⁰We are aware that the high energy here could lead to complications. The circuit picture indicates that there are no significant additional contributions.

to fall into the singularity. In this regime ($t_L - t_{w2} < t_*$),

$$\begin{aligned}\frac{d\mathcal{C}_2(t_L, t_R)}{dt_L} &= -C \frac{du_F^*}{du_L^*} \\ \mathcal{C}_2(t_L, t_R) &= \mathcal{C}_1(t_L = t_{w2}, t_R) - C(t_L - t_{w2})\end{aligned}\tag{A.15}$$

Notice that during this regime when $0 < t_L - t_{w2} < t_*$, as we evolve t_L down, the white hole interior does not expand much due to the collision, and the black hole interior shrinks because the new perturbation is still mild. This will give rise to $2t_*$ switchback. If we ever want to compare black hole geometry with quantum circuit picture, this regime corresponds to the cancellations between forward and backward time evolutions.

As we further push t_L down, once $t_L - t_{w2} > t_*$, we leave the collision region (Figure 28). The white hole starts to expand, while the shrinking of the black hole is significantly delayed due to the second shockwave.

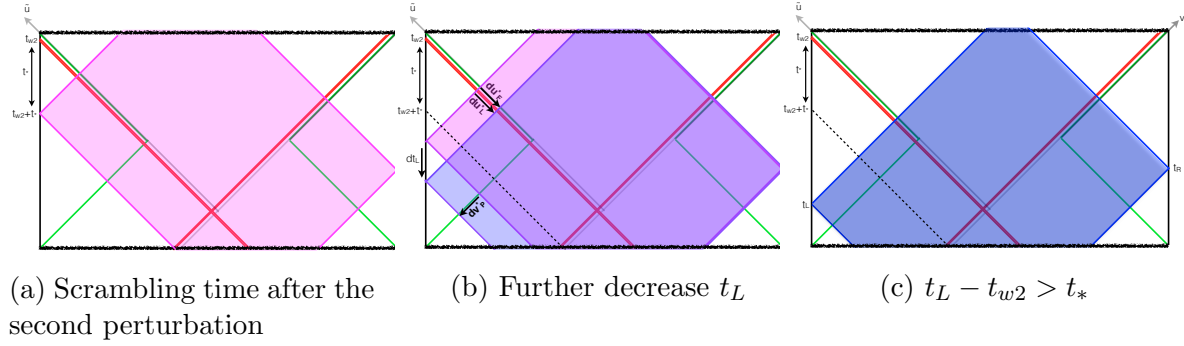


Figure 28: More than scrambling time after the second perturbation

We have

$$\begin{aligned}\frac{d\mathcal{C}_2(t_L, t_R)}{dt_L} &= C \left(1 - \frac{du_F^*}{du_L^*} \right) = C \frac{ce^{\frac{2\pi}{\beta}(t_L - t_{w2} - t_*)}}{1 + ce^{\frac{2\pi}{\beta}(t_L - t_{w2} - t_*)}} \\ \mathcal{C}_2(t_L, t_R) &= \mathcal{C}_2(t_{w2} + t_*, t_R) + C \frac{\beta}{2\pi} \log \left(1 + ce^{\frac{2\pi}{\beta}(t_L - t_{w2} - t_*)} \right) \\ &\quad - C \frac{\beta}{2\pi} \log \left(1 + ce^{\frac{2\pi}{\beta}(t_{w2} + t_* - t_{w2} - t_*)} \right) \\ &= \mathcal{C}_1(t_L = t_{w2}, t_R) + C(t_L - t_{w2} - 2t_*)\end{aligned}\tag{A.16}$$

We notice that in this regime, as we push t_L up, the complexity will decrease, which reflects the fact that the left future horizon is not transparent.

Now we look at the regime when $t_L - t_{w2} > t_*$, $t_{w1} - t_R > t_*$. Plug in $\mathcal{C}_1(t_L = t_{w2}, t_R)$ from (2.11), we have

$$\mathcal{C}_2(t_L, t_R) = \mathcal{C}(|\text{TFD}\rangle) + C[2(t_{w1} - t_*) - t_R + 2(-t_{w2} - t_*) + t_L] \quad (\text{A.17})$$

There are two switchbacks, exactly as in (2.4).

A.3 Breaking spatial translation symmetry: Localized shocks

We can also look at localized perturbations. Their holographic duals are black holes with localized shocks [2]. Assume a planer black hole without perturbations:

$$ds^2 = l^2 \left(-f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 dx^i dx^i \right)$$

where $f(r) = r^2 - r^{3-D}$.

With localized perturbation at an early time, $W_{x_w}(t_w) = e^{iHt_w} W_{x_w} e^{-iHt_w}$, null lines of constant v^* and x again will be pushed forward,¹¹ but this time in a x -dependent way. It was shown in [2] that at large $|x - x_w|$, when crossing the horizon, the null lines will be pushed forward in Kruskal coordinates by the amount

$$\begin{aligned} h(x) &= \frac{\exp \left[\frac{2\pi}{\beta} \left(t_w - \frac{|x - x_w|}{v_B} - t_* \right) \right]}{|x - x_w|^{\frac{D-3}{2}}} \\ &= \exp \left[\frac{2\pi}{\beta} \left(t_w - \frac{|x - x_w|}{v_B} - t_* \right) - \frac{D-3}{2} \log |x - x_w| \right], \end{aligned} \quad (\text{A.18})$$

where $v_B = \sqrt{\frac{(D-1)}{2(D-2)}}$.

This again will cause a mismatch between Eddington-Finkelstein time inside and outside the horizon, in a x -dependent way. If we look at (A.18), when $|x - x_w| \gg 1$, inside the exponential the linear term dominates over the logarithmic term, and the perturbation to spatial location x effectly starts at time $t_w - \frac{|x - x_w|}{v_B}$, so there will be x -dependent time delay. If we look at a fixed value x , the situation is similar to the homogeneous case as discussed in earlier sections once we replace t_w by $t_w - \frac{|x - x_w|}{v_B}$.

Now let's consider the complexity of this perturbed state. We can first consider the contribution from fixed x slice, which we know how it looks like from earlier discussions,

¹¹Contrary to the case with homogeneous perturbations, they are no longer null geodesics.

then do an integration in x . Again, when we do this we implicitly used assumption 2, i.e., locality of tensor network down to AdS scale. Here, a fixed x slice has width l_{AdS} . So far we recovered all time dependence of complexity as discussed in section 2.1.

B Collision energy and time dilation factors

B.1 Raychaudhuri equation

We've seen that we can use a bundle of ingoing light rays to detect the change of Eddington-Finkelstein time across the horizon, and hence detect the time dependence of complexity. Raychaudhuri equation describes the behaviors of a congruence of geodesics [27], in particular, how they respond to stress-energy tensors.

Consider a bundle of ingoing light rays. This bundle can either expand, or shrink as they travel. To quantitatively describe this, consider a small area element A in transverse directions. The expansion rate θ is defined as

$$\frac{dA}{d\tau} = \theta A \quad (\text{B.19})$$

Raychaudhuri equation gives:

$$\frac{d\theta}{d\tau} = -R_{ab}E_0^a E_0^b - 2\sigma^2 + 2\omega^2 - \frac{1}{D-2}\theta^2, \quad (\text{B.20})$$

where ω reflects the rotation of neighboring light rays, σ reflects the distortion of shapes, and θ reflects the area expansion. $E_0 = \partial_\tau$ is the vector field along proper time. From the term $R_{ab}E_0^a E_0^b = 8\pi G_N T_{ab}E_0^a E_0^b$, we see that, the presence of energy causes the bundle of light rays to focus.

With a shell of matter present, the expansion rate θ can suddenly jump, while there will be time dilation $\frac{d\tilde{v}^*}{dv^*}$ across the shell. In case of homogeneous perturbations, radial light rays stay on constant transverse positions as they travel. We have

$$\frac{d\tilde{v}^*}{dv^*} = \frac{\theta}{\bar{\theta}} = \frac{\theta}{\theta + \Delta\theta}, \quad \text{i.e.,} \quad \frac{\Delta\theta}{\theta} = \left(\frac{d\tilde{v}^*}{dv^*}\right)^{-1} - 1. \quad (\text{B.21})$$

With localized perturbations, light rays will be refracted, and the time dilation factors will depend on transverse positions. This will give additional contributions to the expansion

rate:

$$\frac{\Delta\theta}{\theta} = \left[\left(\frac{d\tilde{v}^*}{dv^*} \right)^{-1} - 1 \right] - \left(\frac{\beta v_B}{2\pi} \right)^2 \nabla^2 \left[\left(\frac{d\tilde{v}^*}{dv^*} \right)^{-1} - 1 \right], \quad (\text{B.22})$$

where the Laplacian in (B.22) is in transverse x^i directions.¹²

On the other hand, from (B.19), (B.20), we have:

$$\frac{d\theta}{\theta} = -\frac{1}{D-2} \frac{dA}{A} \left[1 + (D-2) \left(\frac{8\pi G_N T_{ab} E_0^a E_0^b + 2\sigma^2 - 2\omega^2}{\theta^2} \right) \right]. \quad (\text{B.23})$$

In the limit of thin shell, $\frac{\Delta A}{A} \rightarrow 0$, and there won't be singularities in ω and σ since they reflect rotations and separations of neighboring light rays. So the only potential contribution to the sudden jump of expansion rate comes from the stress-energy tensor term $\frac{\Delta A}{A} \frac{8\pi G_N T_{ab} E_0^a E_0^b}{\theta^2}$. This term is proportional to the collision-energy-squared. Near horizons, the large relative boost between the matter and the detecting light beam makes it grow exponentially in time. Combining with (B.21), we see that with homogeneous perturbations, the collision-energy-squared with matter is inversely proportional to the time dilation factor.

With localized perturbations [2], T_{ab} has localized support in transverse spatial directions: $T \propto E e^{\frac{2\pi}{\beta} t_w} \delta^{D-2}(x - x_w)$. The solution of equation (B.22) shows that the time dilation factors $\frac{d\tilde{v}^*}{dv^*}$ can still be proportional to $E e^{\frac{2\pi}{\beta} t_w}$, even at positions where $x \neq x_w$ and T vanishes. Someone falling in there will not encounter matters, but will still be hit hard by gravitational shockwaves. In next section, we'll estimate the collision energy by evaluating the Landau-Lifshitz pseudotensors.

¹²(B.21), (B.22) are exact in the limit of thin matter shell on the horizon. Away from horizon, $T_{ab} E_0^a E_0^b$ won't get too large and matter won't cause significant jump in expansion rate.

B.2 Landau-Lifshitz pseudotensor

To see the energy carried by gravitons, we look at the Landau-Lifshitz pseudotensor [23]:

$$\begin{aligned}
t^{\mu\nu} = \frac{1}{16\pi G_N} & \left[(g^{\mu a} g^{\nu b} - g^{\mu\nu} g^{ab}) (2\Gamma_{ab}^c \Gamma_{cd}^d - \Gamma_{ad}^c \Gamma_{bc}^d - \Gamma_{ac}^c \Gamma_{bd}^d) \right. \\
& + g^{\mu a} g^{bc} (\Gamma_{ad}^\nu \Gamma_{bc}^d + \Gamma_{bc}^\nu \Gamma_{ad}^d - \Gamma_{cd}^\nu \Gamma_{ab}^d - \Gamma_{ab}^\nu \Gamma_{cd}^d) \\
& + g^{\nu a} g^{bc} (\Gamma_{ad}^\mu \Gamma_{bc}^d + \Gamma_{bc}^\mu \Gamma_{ad}^d - \Gamma_{cd}^\mu \Gamma_{ab}^d - \Gamma_{ab}^\mu \Gamma_{cd}^d) \\
& + g^{ab} g^{cd} (\Gamma_{ac}^\mu \Gamma_{bd}^\nu - \Gamma_{ab}^\mu \Gamma_{cd}^\nu) \\
& \left. - 2\Lambda g^{\mu\nu} \right] \tag{B.24}
\end{aligned}$$

Landau-Lifshitz pseudotensor is only a tensor under linear coordinate transformations. To describe the experience of an infalling observer, we evaluate this tensor in the local inertial frame right before the observer crosses the horizon.

Near horizons, assume a near-Rindler metric of the form¹³:

$$\begin{aligned}
ds^2 &= -\frac{4\pi}{\beta}(r - r_H)dt^2 + \frac{dr^2}{\frac{4\pi}{\beta}(r - r_H)} + r^2 dx^i dx^i \\
&= -\frac{\tilde{c}\beta r_H}{\pi} du dv + r_H^2 (1 - \tilde{c}uv)^2 dx^i dx^i \tag{B.25}
\end{aligned}$$

where \tilde{c} is an order one constant depending on the details of the geometry. With localized perturbations, the metric changes to

$$ds^2 = -\frac{\tilde{c}\beta r_H}{\pi} du [dv - \delta(u)h(x)du] + r_H^2 (1 - \tilde{c}uv)^2 dx^i dx^i. \tag{B.26}$$

Kruskal coordinates u, v in (B.25) are not local inertial frames at $u = 0, v = v_0, x = x_0$. We do coordinate transformations:

$$\begin{aligned}
\tilde{u} &= u \\
\tilde{v} &= v - v_0 - \frac{\pi r_H}{\beta} v_0 (1 - \tilde{c}uv_0)^2 (x - x_0)^2 \\
\tilde{x} &= (1 - \tilde{c}uv_0)(x - x_0).
\end{aligned}$$

$\tilde{u}, \tilde{v}, \tilde{x}$ give local initial frame right before an observer crosses the horizon at $\tilde{u} = 0, \tilde{v} =$

¹³This metric is not exactly the Rindler metric. What we need here is an approximation of horizon metric to first order in uv , while in the Rindler metric, the transverse size is a constant.

0, $\tilde{x} = 0$. In these coordinates, the shockwave metric (B.26) becomes¹⁴

$$ds^2 = -\frac{\tilde{c}\beta r_H}{\pi} d\tilde{u} \left[d\tilde{v} - \delta(\tilde{u}) h \left(x_0 + \frac{\tilde{x}}{1 - \tilde{c}\tilde{u}v_0} \right) d\tilde{u} \right] + r_H^2 (1 - 2\tilde{c}\tilde{u}\tilde{v}) d\tilde{x}^2 + \mathcal{O}(\tilde{u}, \tilde{x})^2. \quad (\text{B.27})$$

Evaluating the Landau-Lifshitz pseudotensor (B.24) in this frame, we see that, apart from the cosmological term,

$$t_{uu} = \frac{1}{8\pi G_N} \tilde{c} \delta(u) h(x). \quad (\text{B.28})$$

Assume we also send in a thermal scale quantum from the right hand side. To consider the collision energy, we can look at the quantity appearing in Raychaudhuri equation $-8\pi G_N \frac{\Delta A}{A} \frac{t_{ab} E_0^a E_0^b}{\theta^2}$. From the definition of stress-energy tensor $T^{ab} = p^a \frac{dx^b}{d\tau} \int d\tau \frac{1}{\sqrt{-g}} \delta^{(D)}(x - x(\tau))$, one can verify that¹⁵

$$\frac{E_c^2}{m_p^2} \sim \frac{r_H}{\beta} S^{\frac{D-4}{D-2}} 8\pi G_N \left(\frac{-\Delta A}{A} \right) \frac{t_{ab} E_0^a E_0^b}{\theta^2}.$$

On the other hand, a direct calculation using (B.28) shows that¹⁶

$$-8\pi G_N \frac{\Delta A}{A} \frac{t_{ab} E_0^a E_0^b}{\theta^2} \sim \frac{h(x)}{v} = \left(\frac{d\tilde{v}^*}{dv^*} \right)^{-1} - 1$$

So the collision-energy-squared with gravitons is also inversely proportional to the time dilation factor.

References

- [1] Douglas Stanford and Leonard Susskind. “Complexity and Shock Wave Geometries”. *Phys. Rev.*, D90(12):126007, 2014.
- [2] Daniel A. Roberts, Douglas Stanford, and Leonard Susskind. “Localized shocks”. *JHEP*, 03:051, 2015.

¹⁴Note that here we ignored higher than first order terms in \tilde{u} , \tilde{x} , but we keep $\tilde{u}\tilde{v}$ term, because \tilde{v} is not continuous across the horizon.

¹⁵In fact, one can already see this from the calculations of spherical perturbations (3.12), (B.21), (B.23).

¹⁶The cosmological constant term does not contribute to this.

- [3] Adam R. Brown, Daniel A. Roberts, Leonard Susskind, Brian Swingle, and Ying Zhao. “Holographic Complexity Equals Bulk Action?”. *Phys. Rev. Lett.*, 116(19):191301, 2016.
- [4] Adam R. Brown, Daniel A. Roberts, Leonard Susskind, Brian Swingle, and Ying Zhao. “Complexity, action, and black holes”. *Phys. Rev.*, D93(8):086006, 2016.
- [5] Brian Swingle. Entanglement Renormalization and Holography. *Phys. Rev.*, D86:065007, 2012.
- [6] Brian Swingle. Constructing holographic spacetimes using entanglement renormalization. 2012.
- [7] Thomas Hartman and Juan Maldacena. Time Evolution of Entanglement Entropy from Black Hole Interiors. *JHEP*, 05:014, 2013.
- [8] Leonard Susskind. “The Typical-State Paradox: Diagnosing Horizons with Complexity”. *Fortsch. Phys.*, 64:84–91, 2016.
- [9] Leonard Susskind. Entanglement is not enough. *Fortsch. Phys.*, 64:49–71, 2016.
- [10] Gerard ’t Hooft. On the Quantum Structure of a Black Hole. *Nucl. Phys.*, B256:727–745, 1985.
- [11] Leonard Susskind and Larus Thorlacius. Gedanken experiments involving black holes. *Phys. Rev.*, D49:966–974, 1994.
- [12] Patrick Hayden and John Preskill. “Black holes as mirrors: Quantum information in random subsystems”. *JHEP*, 09:120, 2007.
- [13] Juan Martin Maldacena. Eternal black holes in anti-de Sitter. *JHEP*, 04:021, 2003.
- [14] Yosi Atia and Dorit Aharonov. Fast-forwarding of Hamiltonians and Exponentially Precise Measurements. 2016.
- [15] Stephen H. Shenker and Douglas Stanford. “Black holes and the butterfly effect”. *JHEP*, 03:067, 2014.
- [16] Leonard Susskind and Ying Zhao. “Switchbacks and the Bridge to Nowhere”. 2014.

- [17] Adam R. Brown, Leonard Susskind, and Ying Zhao. Quantum Complexity and Negative Curvature. 2016.
- [18] Stephen H. Shenker and Douglas Stanford. “Multiple Shocks”. *JHEP*, 12:046, 2014.
- [19] Tevian Dray and Gerard ’t Hooft. The Gravitational Shock Wave of a Massless Particle. *Nucl. Phys.*, B253:173–188, 1985.
- [20] Konstadinos Sfetsos. On gravitational shock waves in curved space-times. *Nucl. Phys.*, B436:721–745, 1995.
- [21] Juan Maldacena and Leonard Susskind. Cool horizons for entangled black holes. *Fortsch. Phys.*, 61:781–811, 2013.
- [22] Luis Lehner, Robert C. Myers, Eric Poisson, and Rafael D. Sorkin. Gravitational action with null boundaries. *Phys. Rev.*, D94(8):084046, 2016.
- [23] L. D. Landau and E. M. Lifschits. *The Classical Theory of Fields*, volume Volume 2 of *Course of Theoretical Physics*. Pergamon Press, Oxford, 1975.
- [24] Leonard Susskind, Larus Thorlacius, and John Uglum. The Stretched horizon and black hole complementarity. *Phys. Rev.*, D48:3743–3761, 1993.
- [25] Daniel Harlow and Patrick Hayden. Quantum Computation vs. Firewalls. *JHEP*, 06:085, 2013.
- [26] Daniel Harlow. The Ryu-Takayanagi Formula from Quantum Error Correction. 2016.
- [27] S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2011.